

A Longitudinal Assessment of Coronary Interventional Program Quality

A Report From the American College of Cardiology–National Cardiovascular Data Registry

Lloyd W. Klein, MD, FACC,* Paul Kolm, PhD,† Xin Xu, MS,†
Ronald J. Krone, MD, FACC,‡ H. Vernon Anderson, MD, FACC,§
John S. Rumsfeld, MD, PhD, FACC,|| Ralph G. Brindis, MD, MPH, FACC,¶
William S. Weintraub, MD, FACC†

Chicago, Illinois; Newark, Delaware; St. Louis, Missouri; Houston, Texas; Bethesda, Maryland; and Oakland, California

Objectives This study applied risk adjustment methods to evaluate member institutions of the American College of Cardiology–National Cardiovascular Data Registry with respect to in-hospital mortality in percutaneous coronary intervention patients over a 4-year period to assess variability in risk-adjusted performance measures.

Background Cardiac catheterization laboratories, hospital networks, and third-party payers are interested in assessing the outcomes of percutaneous coronary interventions. Evaluation of outcomes without considering case selection may lead to erroneous conclusions about program quality.

Methods The National Cardiovascular Data Registry database was queried for all percutaneous coronary intervention cases performed between January 1, 2001, and September 30, 2004. Random effects logistic regression was used to develop models of in-hospital mortality and compute an expected mortality rate for each program. The observed mortality rate in each program was divided by the program's predicted rate to obtain the observed/expected (O/E) mortality ratio. Change in the O/E ratio was assessed by a generalized estimating equation approach to repeated measures. An index of variability was calculated by the mean absolute difference between O/E ratios of each pair of years.

Results There were 664,909 interventional procedures performed in 403 National Cardiovascular Data Registry programs from 2001 to 2004. There was no significant systematic change in O/E ratios over the 4-year period, but there was significantly greater variation in O/E ratios associated with lower percutaneous coronary intervention volume programs.

Conclusions Our risk-adjustment models had very good discrimination and were relatively consistent over the study period. There was substantial within-program variation in O/E ratios. This information would provide an indication for a detailed examination of individual programs. (J Am Coll Cardiol Intv 2009;2:136–43) © 2009 by the American College of Cardiology Foundation

From *Rush Medical College, Chicago, Illinois; †Christiana Care Health System, Newark, Delaware; ‡Washington University, St. Louis, Missouri; §University of Texas Health Science Center, Houston, Texas; ||American College of Cardiology, Bethesda, Maryland; and ¶Kaiser Permanente, Oakland, California. E. Magnus Ohman, MD, was the Guest Editor of this report.

Manuscript received May 7, 2008; revised manuscript received September 22, 2008, accepted September 24, 2008.

Cardiac catheterization laboratories, hospital networks, and third-party payers are interested in assessing the outcomes of percutaneous coronary interventions (PCI) and comparing the results objectively across institutions. Evaluating unadjusted clinical outcomes without considering case selection may lead to erroneous conclusions about program quality (1–5). Present trends toward public release of quality information to evaluate programs depending on performance necessitate maintaining, as much as possible, a level playing field for the laboratories involved.

See page 144

Risk-adjusted mortality rate (RAMR) is a standard method of assessing quality of care in the interventional catheterization laboratory. Risk-adjusted mortality rate adjusts on the basis of risk factors that affect, or potentially affect, PCI outcomes and thus compensates to some degree for differences in case mix across institutions. A component of RAMR is the ratio of the observed outcome rate to the expected outcome rate, the latter computed from a regression model of relevant risk factors. The observed to expected (O/E) mortality ratio has been used, along with RAMR, to compare outcomes of various procedures and outcomes across institutions (6–9). Because the expected outcome rate is computed from a model, sampling variability and model discriminatory ability affect the accuracy of the expected rate and thus the O/E ratio and RAMR. In addition, the value of model parameters, or the combination of model parameters, may change over time and consequently affect estimates for individual institutions.

The purpose of this study was to evaluate member institutions of the American College of Cardiology–National Cardiovascular Data Registry (ACC-NCDR) with respect to in-hospital mortality in PCI patients over a 4-year period using risk-adjustment methods. We assessed variation of O/E ratios over the 4-year period and whether O/E ratios differed by institution demographics of location, PCI volume, number of board-certified cardiologists, and teaching or nonteaching status.

Methods

Data registry and selection. The ACC-NCDR is a voluntary national registry that currently receives data from >700 participating hospitals. The basic details of the dataset used here have previously been published (10,11). All data elements are linked to the American Heart Association/American College of Cardiology PCI guidelines. The data collection process used by the ACC-NCDR has been described in detail (11). Data at each participant institution are entered locally into software purchased from vendors certified to accurately acquire and transmit data to the ACC-NCDR. Local institutional PCI programs audited all data for completeness and accuracy. Many local quality

assurance programs are based on the collected data. Additionally, a national audit program sponsored by NCDR has reviewed about 5% of all cases. Only data meeting strict predefined criteria for completeness and accuracy (10,11) are entered into the ACC-NCDR registry and used in this analysis. Each data element is predefined, linked to the American Heart Association/ACC PCI guidelines, and available on the ACC website (12).

Patient population. The NCDR database was queried for all PCI cases performed between January 1, 2001, and September 30, 2004. There were 403 member institutions enrolling patients within this time frame (though not all for all 4 years). These institutions are diverse and widely distributed throughout the country. Member institutions include rural, suburban, and urban centers, both teaching and community hospitals, and institutions of all sizes.

Risk adjustment. The ACC-NCDR previously developed and validated risk-adjustment models for in-hospital mortality in PCI (13,14) using data collected between 1998 and 2000. Using the same set of risk factors (Table 1), models were developed using the current data to calculate expected mortality. Ejection fraction was categorized as “not done,” $\leq 40\%$, and $>40\%$; the latter being the reference category for calculation of odds ratios. Lesion severity was categorized as high, moderate, and low, the latter being the reference category. Acute myocardial infarction was categorized as ST-segment elevated, nonelevated, and no acute myocardial infarction. Smoking history was categorized as current, former, and never smoked, the latter being the reference category. Age was analyzed in 10-year increments and body mass index in 5-kg/m^2 units. Angina class and New York Heart Association functional class were analyzed as ordinal categories. All other risk factors were dichotomous. For this study, we developed separate risk-adjustment models for each year: 2001, 2002, 2003, and 2004.

Random effects logistic regression was used to develop models of in-hospital mortality. The random effects model accounts for patient clustering within institutions. Beginning with the 27 risk factors, a backward elimination strategy was used to develop reduced models. Nested models were compared using the likelihood ratio test. Bootstrap methods were used to validate and calibrate bias-corrected indexes of model performance (15). The C-statistic index (equivalent to the area under a receiver-operator characteristic curve) was calculated to assess model discrimination, the ability of the model to correctly identify patients with respect to in-hospital mortality.

Abbreviations and Acronyms

ACC-NCDR = American College of Cardiology–National Cardiovascular Data Registry

O/E = observed to expected mortality ratio

PCI = percutaneous coronary intervention

RAMR = risk-adjusted mortality rate

Table 1. Risk Factors Assessed in NCDR Risk-Adjustment Models

Family history of CAD	Diabetes
History of renal failure	Previous PCI
Previous MI	Previous CABG
Recent heart failure	Previous valve surgery
Smoking history (current, former, never)	Shock at PCI
Sex (female/male)	Lesions in left main
Race (white/other)	Lesions in proximal LAD
Angina class (1 to 4)	Nonelective PCI
NYHA functional class (0 to 4)	Lesion severity (low, moderate, high)
Chronic lung disease	AMI (none, ST-segment elevated, non-ST-segment elevated)
Cerebrovascular disease	Age (yrs)
Peripheral vascular disease	Ejection fraction (not done, <40%, >40%)
Hypertension	BMI (kg/m ²)
Elevated cholesterol	

AMI = acute myocardial infarction; BMI = body mass index; CABG = coronary artery bypass graft surgery; CAD = coronary artery disease; LAD = left anterior descending; MI = myocardial infarction; NCDR = National Cardiovascular Data Registry; NYHA = New York Heart Association; PCI = percutaneous coronary intervention.

Missing risk factor data were assumed to be missing at random and imputed using Markov Chain Monte Carlo methods (16). A total of 3 imputed datasets were generated for each year. Regression coefficients from the 3 imputed datasets were combined, and a predicted probability of in-hospital mortality was calculated for each patient. Thus, all patients were used in the development of the mortality models.

A logit (natural logarithm of the odds of in-hospital mortality) for each patient was calculated from the logistic regression equations and transformed to a predicted probability of in-hospital death by taking the exponent of the logit and dividing by 1 plus the exponent of the logit. The predicted probabilities were averaged across patients at a participating institution to calculate the predicted mortality at that institution, given its case mix.

The observed mortality rate in each program was divided by the program's predicted rate to obtain the O/E mortality ratio. The O/E ratio yields an index for which a value <1 means that in-hospital mortality was less than expected, and a value >1 means that in-hospital mortality was greater than expected. Programs with the largest O/E ratios would rank the lowest, and programs with smaller O/E ratios would rank higher. Risk-adjusted mortality rate is calculated as the O/E ratio times the overall mortality rate in a given year. In this study, we present results only for the O/E ratio. Bootstrap methods (17) were used to calculate the O/E ratio 95% confidence intervals. Variation in program O/E ratios over the 4 years was assessed by computing the mean absolute difference of O/E ratios, and rankings for each pair of years. A small value would indicate similarity in rankings over the years, whereas a large value would indicate dissimilar rankings.

Generalized estimating equation models for repeated measures were used to compare O/E ratios with respect to number of board-certified cardiologists, PCI volume, location, and teaching status across years. The O/E ratios were right skewed and thus were modeled assuming a gamma distribution. One-way analysis of variance was used to compare absolute differences in rankings averaged over each pair of years for programs with data in all 4 years. SAS version 9.2 (SAS Institute, Cary, North Carolina), and Stata version 10 (Stata Corp., College Station, Texas) were used to develop regression models and compare O/E ratios.

Results

There were 664,909 interventional procedures performed in the 403 programs from 2001 to 2004. The observed mortality rate declined significantly over the 4-year period ($p < 0.001$). Rates were 1.36%, 1.30%, 1.23%, and 1.12% in 2001, 2002, 2003, and 2004, respectively. Complete patient data were available for all 27 risk factors included in the risk models except recent congestive heart failure, New York Heart Association classification, and body mass index. Missing congestive heart failure and New York Heart Association data were <1%, whereas 3% of body mass index data were missing. These variables were imputed as described, and models of in-hospital mortality were obtained by combining the regression coefficients from the 3 imputed datasets.

Table 2 presents the risk-adjustment models developed for each year. Models were relatively consistent from year to year with respect to risk factor composition and impact on mortality. Shock at PCI had the greatest impact on mortality, with odds ratios ranging from 10 to 13. Acute myocardial infarction status, renal failure, left main disease, ejection fraction, nonelective PCI status, and lesion severity were consistent predictors of mortality in all 4 years with odds ratios of 2 or greater. All 4 models had good discrimination with C-statistic indexes of 0.9.

Although there were a total of 403 programs involved over the 4-year period, not all were available for evaluation in every year as some programs may have joined after 2001, some may have left the program before 2004, or possibly, some may not have reported data in every year. Of the 403 programs, 45% had data for all 4 years, 21% for 3 of the 4 years, 24% for 2 years, and 10% had data for only 1 of the years. For 2001, there were 228 programs with data for evaluation. A total of 41 (18%) of these programs had O/E ratios significantly >1 and ranged from 1.30 to 5.96. There were 285 programs evaluated in 2002, and 69 (24%) of these had O/E ratios significantly >1, ranging from 1.20 to 4.54. For 2003, 355 programs were evaluated, and 90 (25%) of these had O/E ratios significantly >1, ranging from 1.13 to 7.71. For 2004, 339 programs were evaluated, and 54 (16%)

Table 2. Predictive Models of In-Hospital Mortality by Year

Risk Factor Type	Risk Factor	Category	2001	2002	2003	2004
Demographic/history	Age (10 yrs)		1.55	1.59	1.58	1.57
	Male sex		1.28	1.31	1.36	1.61
	Family history of CAD		0.80	0.88	0.77	0.85
	Previous PCI		0.81	0.89	0.85	
	Previous MI		1.20			
Comorbidity	Acute MI	ST-segment elevated	3.30	3.70	4.17	5.08
		Non-ST-segment elevated	2.17	2.27	2.08	2.88
	Renal failure		2.69	2.37	2.40	2.41
	Heart failure			0.61	0.69	0.68
	Cerebrovascular disease			1.40	1.32	1.16
	Chronic lung disease		1.51	1.54	1.39	1.41
	Peripheral vascular disease	1.36	1.39	1.33	1.57	
	Diabetes		1.30	1.30	1.29	1.26
	Elevated cholesterol		0.71	0.62	0.66	0.65
Disease severity	Shock at PCI		13.48	10.86	11.02	13.40
	Left main disease		2.23	2.21	2.33	2.12
	Ejection fraction	Not done	2.37	2.17	2.33	2.26
		≤40%	2.00	2.27	2.16	2.14
	Proximal LAD		1.24	1.16	1.15	1.29
	Nonelective PCI		2.81	2.41	2.48	2.19
	NYHA functional class		1.25	1.44	1.38	1.32
Lesion severity	AHA classification	High	3.13	2.61	2.15	2.00
		Moderate	1.48	1.36	1.20	1.17
Number of programs			228	285	355	339
Data are odds ratios. All are significantly different from 1 except moderate American Heart Association (AHA) classification in 2004 (1.17). History of heart failure is associated with increased risk of mortality, but in combination with the other risk factors in the model, heart failure odds ratios were significantly <1. Abbreviations as in Table 1.						

of these had O/E ratios significantly >1, ranging from 1.22 to 5.77.

There was considerable variability of O/E ratios among the various programs over the 4-year period. For presentation, we chose the 20 lowest ranked (largest O/E ratios) programs in each year (Table 3). Of the 403 programs, 339 (84%) were never among the 20 lowest-ranked programs in the 4-year period. There were a total of 64 programs that were among the 20 lowest-ranked programs over the 4-year period. Of the 64, 1 program had data for only 1 year, 9 programs had data for 2 years, 18 had data for 3 years, and the remaining 36 had data for all 4 years. Of the 64 programs, 53 (83%) appeared among the 20 lowest-ranked only once in the 4-year period. Eight appeared in 2 of the 4 years, 1 appeared in 3 of the 4 years, and 2 were ranked among the 20 lowest all 4 years.

There were a total of 180 programs (45%) for which O/E ratios could be computed in all 4 years. Using these data, the top panels of Figure 1 plot O/E ratios for the 20 lowest, 20 middle, and 20 highest-ranked programs in 2001 and corresponding O/E ratios in subsequent years. Generally, there was greater variability of O/E ratios over the years in the lowest-ranked programs, as well as the highest-ranked

programs, when compared with middle-ranked programs. The lower panels highlight 2 programs from each group in the top panels. Programs 2, 4, 45, 54, and 29 are identified by the same number in Table 3 and illustrate yearly variation in rankings.

Analysis of O/E ratios across years indicated no significant relationship with location ($p = 0.46$), teaching status ($p = 0.20$), number of board-certified cardiologists ($p = 0.22$), or PCI volume ($p = 0.24$). Analyzing only those programs with 4 years of data also indicated no significant relationships ($p = 0.47, 0.82, 0.95$, and 0.18 for location, teaching status, cardiologists, and PCI volume, respectively). Comparing mean absolute difference in rankings for each pair of years indicated no significant differences between locations ($p = 0.86$), teaching status ($p = 0.51$), and number of board-certified cardiologists ($p = 0.37$). Lower-volume programs (<400 per year) had significantly greater differences in rankings over the years than medium- (400 to 800 per year) and high- (>800 per year) volume programs ($p < 0.001$ for both), likely reflecting sampling variability. Lower-volume programs had a mean absolute difference of 53 ± 20 places in rankings over the 4-year period compared with 40 ± 19 and 33 ± 16 for medium- and high-volume programs, respectively. The same

Table 3. 20 Lowest-Ranked NCDR Programs, 2001 to 2004

2001		2002		2003		2004	
PID	O/E (95% CI)	PID	O/E (95% CI)	PID	O/E (95% CI)	PID	O/E (95% CI)
1	5.96 (4.37–9.18)	2	4.54 (4.04–5.13)	38	7.71 (4.43–14.61)	51	5.77 (4.41–7.80)
2	4.06 (3.09–5.21)	9	3.83 (3.04–4.70)	1	5.96 (4.36–8.70)	43	4.73 (3.91–5.75)
3	3.74 (2.85–4.81)	1	3.68 (3.14–4.38)	32	5.28 (4.63–6.03)	52	4.72 (2.94–8.21)
4	3.49 (2.35–5.38)	21	3.63 (2.31–5.80)	33	4.81 (3.55–6.66)	2	4.20 (3.62–4.88)
5	3.43 (2.54–4.80)	22	3.47 (1.84–7.02)	28	4.32 (3.40–5.46)	53	3.69 (2.90–4.79)
6	2.83 (1.65–5.97)	23	3.22 (2.07–5.03)	39	4.13 (3.60–4.75)	54	3.24 (2.43–4.42)
7	2.82 (2.12–3.94)	24	3.11 (1.95–6.24)	40	3.68 (2.25–7.28)	9	3.23 (2.69–3.87)
8	2.66 (1.77–4.58)	25	3.06 (2.30–4.15)	41	3.35 (2.30–5.17)	55	2.38 (1.58–3.68)
9	2.50 (2.01–3.06)	26	2.97 (2.21–3.81)	42	3.04 (1.91–5.47)	56	2.37 (1.66–3.44)
10	2.21 (1.74–2.85)	27	2.97 (1.65–5.73)	43	2.92 (2.29–3.78)	57	2.34 (1.45–3.80)
11	2.16 (1.50–3.41)	28	2.85 (2.30–3.62)	2	2.83 (2.18–3.51)	58	2.31 (1.67–3.57)
12	2.11 (1.50–3.21)	29	2.81 (2.05–4.02)	44	2.81 (1.75–5.25)	38	2.25 (1.69–3.10)
13	2.08 (1.26–3.79)	30	2.76 (2.07–3.80)	45	2.79 (2.18–3.58)	59	2.24 (1.66–2.90)
14	1.98 (1.45–2.91)	31	2.47 (2.18–2.86)	46	2.71 (2.19–3.35)	60	2.16 (1.69–2.81)
15	1.96 (1.59–2.51)	32	2.41 (1.93–3.02)	9	2.66 (2.27–3.10)	61	2.13 (1.60–2.91)
16	1.86 (1.39–2.59)	33	2.37 (2.08–2.71)	47	2.61 (1.87–3.81)	16	2.07 (1.49–2.88)
17	1.86 (1.36–2.82)	34	2.32 (1.41–4.05)	35	2.46 (1.68–3.43)	62	2.02 (1.53–2.79)
18	1.84 (1.39–2.44)	35	2.27 (1.69–3.02)	48	2.41 (1.99–2.97)	63	1.94 (1.30–3.04)
19	1.82 (1.42–2.35)	36	2.21 (1.55–3.28)	49	2.37 (1.93–2.90)	64	1.91 (1.28–2.99)
20	1.81 (1.54–2.14)	37	2.14 (1.10–4.61)	50	2.34 (1.60–3.71)	13	1.87 (1.15–3.49)

Data are observed to expected (O/E) mortality ratios and 95% confidence intervals (CI). Programs are numbered consecutively from 1 to 64 with program IDs (PID) repeated for programs appearing in more than 1 year. The IDs in **bold** are those appearing in the 20 lowest-ranked programs 2 or more years.

Abbreviations as in Table 1.

findings were obtained for mean absolute difference in O/E ratios (data not shown).

Discussion

Use of outcomes to assess coronary interventional program quality has become a common practice and is used in public domain reports such as the California CABG Mortality Reporting Program (18) and the New York Cardiac Surgery Reporting System (19). Such reports employ risk adjustment methods, because observed outcome rates are not adjusted for case mix and therefore are not usually appropriate for comparison of programs or individuals. Reporting of risk-adjustment outcomes for individual institutions or physicians is regarded as the more appropriate measure of quality (20–22).

A number of studies have been done to assess the validity and use of risk-adjustment methods (1,7,8). Because risk-adjustment models involve sampling variability in estimating predicted probabilities of the outcome, program size, as well as number of events, will affect the confidence intervals of O/E ratios (or RAMR). In particular, smaller programs may have wider confidence intervals that include 1 for O/E (or 0 for RAMR) even though the ratio may be large. This would be the case whether confidence intervals were obtained by bootstrap or normal theory methods.

The discriminatory ability of the model will affect the accuracy of predicted outcomes. If model discrimination were low, then prediction accuracy may not be sufficient to use for program evaluation. Model discrimination for in-hospital mortality in our study was quite high (0.9), so there would be good confidence in the estimated expected probabilities for calculation of the O/E ratio. On the other hand, model discrimination for the outcome of unplanned coronary artery bypass graft in our data was much lower, about 0.7 (data not shown). Thus, use of unplanned coronary artery bypass graft as a measure of program quality would be less valid than in-hospital mortality. Setting a value for acceptable discrimination would be arbitrary, but it would be important for program evaluators to understand that an index of 0.5 indicates no discrimination and the closer the index to 1.0, the better the discrimination.

Our study evaluated NCDR participating programs over a 4-year period from 2001 to 2004 with respect to in-hospital mortality. An O/E ratio and 95% confidence interval was obtained for each program in each year. There were, on average, 40 programs each year that had fewer than 100 patients. However, over the 4-year period, only 9 programs with fewer than 100 patients had O/E ratios of 1.5 or greater and 95% confidence intervals that included 1. Thus, most NCDR programs had sufficient patient volume so that a decision whether their O/E ratio was significantly

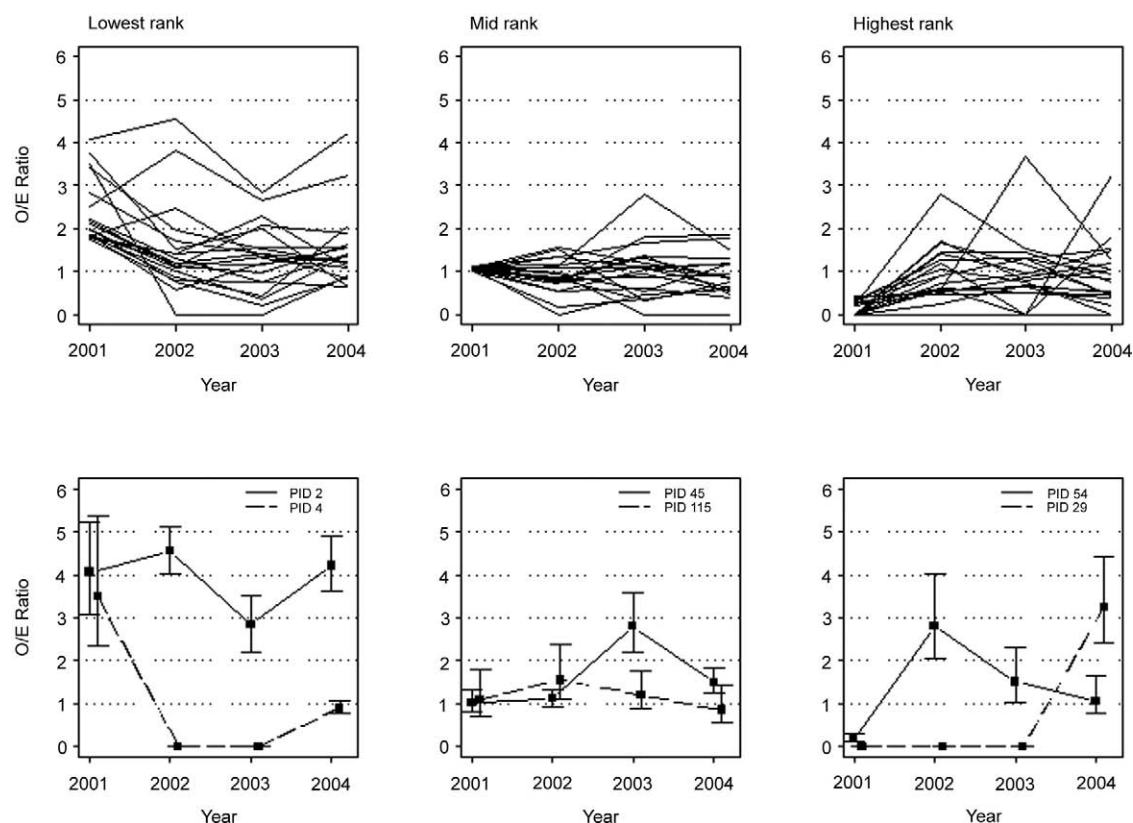


Figure 1. O/E Ratios Over 4-Year Period for 20 Lowest-, 20 Middle-, and 20 Highest-Ranked Programs in 2001

Observed to expected (O/E) mortality ratios over 4-year period for 20 lowest, 20 middle, and 20 highest ranked programs in 2001. PID = program ID.

different from 1 was not an issue. Program evaluation of outcomes related to patient volume itself has been studied but has been questioned as an appropriate measure of quality (23–26). Our results indicated that there was no difference with respect to O/E ratios between programs with differing patient volumes, but the variability in year-to-year rankings and O/E ratios was greater in lower volume programs.

Our data suggest that a few NCDR participant programs were habitually problematic and others fell in and out of the lowest O/E ratio rankings. This may reflect real changes in programmatic quality due to local quality assurance programs that used their data to identify systematic problems and fix them. However, chance also affects placement. If a longitudinal assessment is considered, as in our study, then in a 1-year period, the probability of a program appearing by chance among the 20 lowest-ranked of 400 programs is 1 in 20 (20 of 400). The probability of being ranked, by chance, in the lowest 20 in 2 of the years is 6 in 400 ($4! / [(4-2)! \cdot 2!] \cdot 1/20 \cdot 1/20$). For 3 of the years, the probability decreases to 4 in 8,000, and for all 4 years, 1 in 160,000. For a larger number of programs, the probabilities would be much smaller.

A ranking among the lowest in 1 year may not be adequate to judge a program's performance, but being ranked lowest in more than 1 of the years by chance is small and would likely be a cause for concern. The variation in individual O/E ratios (and rankings) from year to year indicates that "program quality," at least in terms of mortality, is not as readily apparent as it might if only 1 year's data were analyzed. For example, program #2, with large O/E ratios all 4 years, would probably want to evaluate their program in more depth to determine whether there was a systematic problem at the hospital level, or whether a particular operator required further scrutiny. However, program #29 may also want to evaluate their program to determine the reason for the large O/E ratio after 3 consecutive years of no deaths (Fig. 1).

We found no evidence of systematic change in O/E ratios over the 4-year period: that is, there was no overall trend, and none related to program demographics of PCI volume, number of board-certified cardiologists, or teaching status. There was considerable individual program variation in O/E ratios, and this was related to PCI volume with larger variation associated with smaller volume programs.

Study limitations. The primary limitation of this study, or any other where reporting is voluntary, is that member institutions were not necessarily the same from year to year. This may affect the models developed in each year and thus impact expected mortality differently from year to year. However, in a large data registry, this impact may be minimal, and our results suggest this is the case as the models were relatively similar from year to year and consistent with respect to discrimination. The prediction model itself is influenced by case selection (8) as well as advances in technique and adjunctive pharmacology. We developed models for each year, but applying the 2001 model to subsequent years resulted in some loss of accuracy, so the model would need to be recalibrated (7). At some point, it would likely be necessary to update the model.

Calculation of a meaningful institutional O/E ratio is dependent on adequate patient volume. For example, the death of 1 patient or 2 patients in a low-volume site that has 10 cases in a given year will yield a high mortality rate that may not accurately reflect the program's quality. There were a small number of NCDR participant low-volume sites, as assessed in relative terms. However, the distribution of volume in NCDR is characteristic of the practice of medicine in the U.S., and NCDR participant institutions represent almost 30% of all PCI programs. It is possible that mainly institutions with a strong commitment to quality participate in NCDR, and thus, the NCDR may not reflect all institutions in their volume category.

It is also possible that certain operators might "game" the system: that is, recognizing that their data is being collected, they may purposely exaggerate the presence of high risk variables that will lead to higher estimated mortality scores. However, there is no objective evidence of such a practice, and there is a robust external audit system in place designed to minimize this concern.

Conclusions

Third-party payers, including insurance carriers and centers for Medicare and Medicaid services, are demanding an objective method to assess program quality of individual PCI programs. Risk-adjustment methods provide a way of making a comparison of outcomes more equitable by adjusting for case mix. In particular, the O/E ratio is an easily interpreted metric that indicates a program's standing with respect to worse-than-expected or better-than-expected outcomes. In our study, the risk-adjustment models used to calculate the expected outcome had very good discrimination. Over a 4-year period, there was no systematic change in O/E ratios, but there was substantial individual program variability. Variation in O/E ratios over some time frame could be provided for individual programs in the form of "quality control" plots (as in the lower panels of Fig. 1), and comparison with other programs by rankings, or percentiles

of rankings, based on O/E ratios. Although the O/E ratio has a straightforward interpretation, understanding how it is derived, the effects of sampling variability, and model accuracy on its derivation would help institutions to use it as an indication for more detailed examination of their program rather than as an absolute number that defines the program as simply "high" or "low" quality.

Reprint requests and correspondence: Dr. Lloyd W. Klein, Director, Clinical Cardiology Associates, Gottlieb Memorial Hospital, 675 West North Avenue, Melrose Park, Illinois 60160. E-mail: lloydklein@comcast.net.

REFERENCES

1. Moscucci M, Eagle KA, Share D, et al. Public reporting and case selection for percutaneous coronary interventions. An analysis from two large multicenter percutaneous coronary intervention databases. *J Am Coll Cardiol* 2005;45:1759–65.
2. Brindis RG, Dehmer GJ. Continuous quality improvement in the cardiac catheterization laboratory. Are the benefits worth the cost and effort? *Circulation* 2006;113:767–70.
3. Epstein AJ, Rathshore SS, Krumholz HM, Volpp KGM. Volume-based referral for cardiovascular procedures in the United States: a cross-sectional regression analysis. *BMC Health Serv Res* 2005;5:42–50.
4. Epstein AJ, Rathshore SS, Volpp KGM, Krumholz HM. Hospital percutaneous coronary intervention volume and patient mortality, 1998–2000. Does the evidence support current procedure volume minimums? *J Am Coll Cardiol* 2004;43:1755–62.
5. McGrath PD, Wennberg DE, Dickens JD, et al. Relation between operator and hospital volume outcomes following percutaneous coronary interventions in the era of the coronary stent. *JAMA* 2000;284:3139–44.
6. Racz MJ, Hannan EL, Isom OW, et al. A comparison of short- and long-term outcomes after off-pump and on-pump coronary artery bypass graft surgery with sternotomy. *J Am Coll Cardiol* 2004;43:557–64.
7. Peterson ED, DeLong ER, Muhlbaier LH, et al. Challenges in comparing risk-adjusted bypass surgery mortality results: results from the Cooperative Cardiovascular Project. *J Am Coll Cardiol* 2000;36:2174–84.
8. Glance LG, Osler T, Shinozaki T. Effect of varying case mix on the standardized mortality ratio and W statistic: a simulation study. *Chest* 2000;117:1112–7.
9. Cleveland JC Jr., Shroyer ALW, Chen AY, Peterson E, Grover FL. Off-pump coronary artery bypass grafting decreases risk-adjusted mortality and morbidity. *Ann Thorac Surg* 2001;72:1282–9.
10. Anderson HV, Shaw RE, Brindis RG, et al., for the ACC-NCDR. A contemporary overview of percutaneous coronary interventions. *J Am Coll Cardiol* 2002;39:1097–101.
11. Weintraub WS, McKay CR, Riner RN, et al., for the American College of Cardiology Database Committee. The American College of Cardiology National Database: progress and challenges. *J Am Coll Cardiol* 1997;29:459–65.
12. ACC National Cardiovascular Data Registry. Cardiac Cath Lab Module, version 2.0c [pdf]. Available at: www.accncdr.com/WebNCDR/NCDRDocuments/DataDictDefsOnlyv20c.pdf. Accessed June 7, 2007.
13. Shaw RE, Anderson HV, Brindis RG, et al., for the ACC-NCDR. Development of a risk adjustment mortality model using the American College of Cardiology–National Cardiovascular Data Registry (ACC-NCDR) experience: 1998–2000. *J Am Coll Cardiol* 2002;39:1104–12.
14. Shaw RE, Anderson HV, Brindis RG, et al., on behalf of the ACC-NCDR. Updated risk adjustment mortality model using the

- complete 1.1 dataset from the American College of Cardiology Data Registry (ACC-NCDR). *J Invasive Cardiol* 2003;15:578–80.
15. Harrell FE Jr. *Regression Modeling Strategies*. New York, NY: Springer, 2001.
 16. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med* 1999; 18:681–94.
 17. Efron B. *The Jackknife, the Bootstrap and Other Resampling Plans*. Montpelier, VT: Capital City Press, 1982.
 18. California CABG Outcomes Reporting Program (CCORP). Data Abstractor Training Handbook Version 2.0. Available at: http://www.oshpd.cahwnet.gov/HID/SubmitData/CCORP_CABG/2006AbstractTrain.pdf. Accessed June 7, 2007.
 19. New York State Department of Health. Adult Cardiac Surgery in New York State: 2002–2004 [pdf]. June 2006. Available at: http://www.health.state.ny.us/diseases/cardiovascular/heart_disease/docs/cabg_2002-2004.pdf. Accessed June 7, 2007.
 20. Klein LW, Shaw RE, Krone RJ, et al., for the ACC-NCDR. Mortality after emergent percutaneous coronary intervention in cardiogenic shock secondary to acute myocardial infarction and usefulness of a mortality prediction model. *Am J Cardiol* 2005;96:35–41.
 21. Finalyson SRG. Delivering quality to patients. *JAMA* 2006;296: 2026–7.
 22. Go V, Calvin JE, Klein LW. Independent validation of the American College of Cardiology–National Cardiovascular Data Registry risk model to predict complications of percutaneous coronary interventions (abstr). *Circulation* 2002;106:II717.
 23. Vakili BA, Kaplan R, Brown DL. Volume-outcome relation for physicians and hospitals performing angioplasty for acute myocardial infarction in New York State. *Circulation* 2001;104:2171–9.
 24. Brindis RG, Weintraub WS, Dudley RA. Volume as a surrogate for percutaneous coronary intervention quality: is this the right measuring stick? *Am Heart J* 2003;146:932–4.
 25. Brown DL. Analysis of the institutional volume-outcome relationship for balloon angioplasty and stenting in the stent era in California. *Am Heart J* 2003;146:1071–6.
 26. Kuntz RE, Normand SLT. Measuring percutaneous coronary intervention quality by simple case volume. *Circulation* 2005;112: 1088–91.
-
- Key Words:** percutaneous coronary intervention outcomes ■ risk-adjusted models ■ program quality.